

# DOCUMENT RESUME

ED 193 580

CG 014 741

AUTHOR Stevens, George E.  
 TITLE Teaching by Whose Objectives? The View of Students and Teachers.  
 PUB DATE [78]  
 NOTE 22p.  
 EDRS PRICE MF01/PC01 Plus Postage.  
 DESCRIPTORS \*College Faculty; College Students; \*Evaluation Criteria; \*Faculty Evaluation; Feedback; Higher Education; Performance; Student Attitudes; \*Student Evaluation of Teacher Performance; Student Teacher Relationship; \*Teacher Characteristics; \*Teacher Effectiveness; Teacher Improvement

## ABSTRACT

Student and faculty respondents were asked to rate the importance of selected teacher traits in terms of student learning, and to make judgments concerning the value and use of student evaluations of teachers. The results indicate that students and faculty differ somewhat in the importance assigned to some teacher traits, such as "instructor indicates what is important for examinations" (valued less highly by faculty). Many faculty members believed that students give lower ratings to more demanding professors; students disagreed with this perception. Students and teachers agreed that student evaluations should not be used for personnel decisions. Regarding the central issue of whether students and faculty use different criteria in evaluating teacher effectiveness, statistically significant differences were found between the two groups. Given these differences, student-evaluation instruments might best be used only for the purpose of giving teachers feedback. (Author/CS)

\*\*\*\*\*  
 \* Reproductions supplied by EDRS are the best that can be made \*  
 \* from the original document. \*  
 \*\*\*\*\*

ED193580

TEACHING BY WHOSE OBJECTIVES? THE  
VIEW OF STUDENTS AND TEACHERS

George E. Stevens  
College of Business Administration  
Arizona State University

Running Head: Teaching by Whose

Address correspondence to:

Dr. George E. Stevens  
Department of Management  
College of Business Administration  
Arizona State University  
Tempe, Arizona 85281  
(602) 965-2227 or 965-3431

U.S. DEPARTMENT OF HEALTH,  
EDUCATION & WELFARE  
NATIONAL INSTITUTE OF  
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS  
MATERIAL HAS BEEN GRANTED BY

*George Stevens*

TO THE EDUCATIONAL RESOURCES  
INFORMATION CENTER (ERIC)."

CG 014741

The most common performance appraisal device used in academia today is the student evaluation of teacher effectiveness. Typically, the student, near the completion of a course, completes a standardized evaluation form for the teachers of the class in which s(he) is enrolled. Evidence of the widespread use of teacher evaluations has been documented by a number of researchers (e.g., Lein & Merz, 1978; Peterson, Kerin & Martin, 1978). For example, in a study designed to identify the methods used for evaluation of business faculty, Lein and Merz (1978) received responses from 374 business schools. Although respondents indicated that they used various combinations of methods in evaluating business faculty, over 70 percent of the schools used some form of teacher evaluation by students. Not only are more schools using this method of assessing teacher effectiveness, many also use the results to make administrative decisions (e.g., faculty retention, promotion, salary and tenure).

As usage of student evaluations of teacher performance increases, so has the amount of literature reporting the uses and abuses of these devices (see e.g., Miller, 1978; and Miller, Brokaw, & Shaaban, 1977). It is evident that there are both proponents and opponents of the use of student evaluations of teachers as input into personnel decisions. Most faculty members agree that these evaluations have value if used for faculty development purposes but are leery of their usage for other purposes. One reason for this concern is the many reliability and validity issues related to teacher evaluations; issues which have been investigated by a number of researchers. Researchers have

discovered, for example, that many who construct such ratings are not sufficiently qualified to do so (Costin, Greenough & Menges, 1971). Furthermore, when colleague and supervisor ratings of teacher effectiveness were also obtained, low correlations were found between colleague or supervisor ratings and student ratings. Other researchers (e.g., Rodin & Rodin, 1972) conclude that students are not able to judge teaching effectiveness.

Many variables have been identified which influence student perceptions of teacher effectiveness. In many cases, either the teacher cannot control the variables or the variables may be difficult to measure. Studies undertaken include those examining student attributes such as student achievement (Banziger & Smith, (Note 2); Costin, et al., 1971), achievement factors (Banziger & Smith, (Note 2)); personality traits (Warren & O'Connell, 1978); and sex of student (Wilson & Doyle, 1976). Also, a number of other variables have been examined including leader behavior or style (Swanson, 1975; Kinicki & Schriesheim, 1978; Baba & Ace, (Note 1)), type of course, i.e., required vs. elective (Miller, 1978; Miller, et al., 1977), teacher demands (Sullivan and Skanes, 1974), class size (Miller, 1978; Miller, et al., 1977), sex of teacher (Elmore & LaPointe, 1975; Wilson & Doyle, 1976), and teacher personality (Elmore & LaPointe, 1975; Witty, 1947). Although full discussion of these issues is beyond the scope of this paper, the interested reader is directed to see reviews such as those of Costin, Greenough, & Menges (1971) or Sullivan & Skanes (1974).

Although literature on the subject of student evaluations has proliferated, few researchers have studied the reactions of students and teachers to their usage. Attitudinal studies, when undertaken, tend to focus upon a specific course and professor (see e.g., Schwab, 1977). The review by Costin, et al. (1971) and a study by Penfield (1978) are notable because the researchers elicit students' reaction to the student evaluation process. However, no studies were found which report teacher reactions to the process. A third study, done by Baum and Brown (Note 3) is also noteworthy. These researchers investigated the extent to which faculty and students embrace different criteria in evaluating teaching effectiveness. Baum and Brown found that students and faculty placed significantly different weights on eight of the ten traits examined. The increasing use of student evaluations in general and the use of these results for administrative decisions, and the paucity of research in the area of student and faculty reaction to the usage of these instruments indicates a research topic in need of further exploration. In addition, research in the area of student evaluation, particularly in colleges of business, has been a neglected issue. A majority of the work in this area has been performed by researchers in colleges of education. One notable exception is a study by Weaver-Wilson (Note 4) in which she explored the possibility of differences in business student evaluations relative to the student ratings from other colleges.

Additional research into student and faculty perceptions of selected teaching traits and the process of teacher evaluation would provide valuable insights into the dynamics of teacher performance appraisal and fill an important void in the current literature. If, as Cummings and Schwab's (1973) findings suggest, supervisors and subordinates tend to not agree as to the dimensions of the subordinate's job, the relative importance of the job dimensions, and the individual's achievements on these dimensions, then the meaningfulness and validity of performance evaluation data is open to question. For example, Maier, Hoffman, Hooven, & Read (1961) found substantial disagreement between manager-subordinate pairs on the subordinate's job duties, job requirements, and the obstacles that hinder the subordinate. Many people assume that students are in the best position to judge a teacher's performance. The results of the studies cited indicate a need to test this assumption. Given the knowledge that such disagreements have taken place in industrial settings, it is quite likely that similar disagreement exist between teacher ratees and student raters in academia.

The relationship between the study by Maier et al. (7) and the present study is that of determining if raters are using different criteria than ratees in evaluating teaching effectiveness and whether students and teachers perceive the student evaluation process differently. If differences do exist, then these differences may have a dramatic impact upon student evaluation results. In fact,

if different criteria are used and if the student evaluation process is perceived differently, it may be inappropriate to use these results as major inputs for administrative decisions. It is believed, therefore, that additional research into student and faculty perceptions of teaching effectiveness and the student evaluation process would add to our knowledge and fill an important void in our understanding. The present study takes a small step in this direction by asking those most affected by the process about the importance of selected teaching traits and the use of student evaluations. Specifically, ten teacher traits and behaviors were identified. Student and faculty respondents were asked to rate the importance of these factors in terms of student learning. Additionally, these individuals responded to a 15-item scale concerning the value and use of teacher evaluations. Two specific hypotheses were postulated:

#### Hypotheses

1. There will be significant differences between faculty and students on their ratings of the importance of selected teacher traits.
2. There will be significant differences between faculty and students in their perceptions of the value and appropriate use of student evaluations.

Method

A total of 226 faculty members and 572 students at a College of Business in a southwestern university participated in the study. All students were undergraduate business students enrolled in various management courses. Faculty respondents were full time teachers in the College. The average faculty respondent is approximately 40 years of age, has attained the rank of assistant professor, and serves on 2 committees. Seventy-two percent of these respondents have published at least one journal article and thirty-nine percent have written a book. On average, these respondents have 10½ years of teaching experience. The average student who participated in the survey is approximately 23 years old, is a junior who has a "B" average and works about 17 hours a week. All subjects responded to the survey voluntarily, and all responses were kept anonymous. A convenience sample of students was used. Two questionnaires were administered to the students during regular class time while faculty member questionnaires were placed in campus mail-boxes. Approximately 65 percent of the faculty returned completed questionnaires. Faculty respondents and nonrespondents did not differ in any significant manner. The questionnaires included a cover letter assuring the respondent of complete confidentiality. First the respondent completed a section of the questionnaire which asked for demographic information, e.g., respondent age and sex. Students also gave information about course hours taken, academic major, hours employed per week, etc.

Incomplete questionnaires were obtained from fifteen students and nine faculty members. These questionnaires were dropped from the analysis, leaving a set of usable responses from 559 students and 217 faculty.

### Instruments

Teaching Effectiveness. Each respondent completed a short questionnaire listing ten teaching traits and behaviors that are commonly exhibited in a classroom situation, particularly when a faculty member utilizes a lecture format in presenting material. The instrument was developed by Baum & Brown (Note 3), and employed in their study of student and faculty perceptions of teaching effectiveness. For each of the ten factors, respondents rated the importance of the item (in terms of student learning) on a seven point Likert-type scale ("1" = not at all important; "7" = extremely important).

Student Evaluation. Respondent reactions to the use of student evaluations as a measure of teaching effectiveness were assessed using a scale developed by the author. The instrument consists of 15 statements about what it is that student evaluations measure and how these data are used or should be utilized. Respondents were asked to indicate how strongly they agreed or disagreed with each statement using 7-point Likert scales ("1" = strongly disagree; "7" = strongly agree).

### Results

Hypothesis 1 stated that faculty and student responses regarding the importance of selected teacher traits would be significantly different. Mean scores of both groups were calculated and compared for each of the ten traits. Comparison of these means with t-tests reveal that for eight items, the difference between the response means is significant at  $p < .01$ ; one of the two remaining is significant to the  $p < .05$  level. Data relevant to Hypothesis 1 (which discusses teacher traits) are presented in Table 1.

-----  
Table 1 Goes About Here  
-----

As can be seen, not only are the differences significant, but it is also clear that faculty and students rank the items differently in terms of importance. Specifically, traits 3, 8, and 9 show the largest differences. Faculty rated these traits 8th, 6th, and 5th respectively, while students rated them 2nd, 9th, and 9th (these last two were tied). The most extreme difference in ratings of importance was for the term "instructor indicates what is important for examinations." Faculty valued this less highly than students ( $t = 8.87$ ). Further examination of these data, however, indicate areas of agreement by faculty and students; for example, trait 6 is rated similarly by both groups. Additional evidence of agreement is provided when the traits are ranked in each group. The Spearman rank correlation between the faculty and student rankings of the selected teacher traits is  $+ .482$ ,  $p < .05$ . This positive correlation indicates some general agreement between faculty and students.

Hypothesis 2 stated that faculty and student perceptions of the values and usage of student evaluations would differ significantly. Mean scores of both groups were calculated and compared for each of the 15 items on the student evaluation questionnaire. Comparison of these means by  $t$ -tests reveals significant differences in the two groups' responses for thirteen of the fifteen items. Twelve of the items are significant at the  $p < .001$  level and two are significant at the five percent level. These results are presented in Table 2. Again, it should be noted that despite the statistically significant differences which are shown in Table 2, there is, practically speaking, some agreement between the faculty and students on many of the items.

-----  
Table 2 Goes About Here  
-----

As can be seen in Table 2, students and faculty are undecided as to whether teacher evaluations actually measure teacher effectiveness. (The mean differences for this particular item, however, are significant). Although students disagree that they use teacher evaluations to punish professors, the teachers responding are undecided. The responses on this item differ significantly  $t(774) = 6.25, p < .001$ . While faculty respondents were not too sure that students are qualified to appraise their performance, students tended to agree slightly that they were qualified to do so ( $t = 3.98, p < .001$ ). Both groups disagreed with the statement that evaluations served no useful

purpose, but faculty were significantly stronger in their disagreement ( $t = 3.20, p < .001$ ). Although mean responses were not statistically significant, it should be noted that both groups were strongly in favor of the continued use of teacher evaluations.

In general, both groups were in agreement or disagreement with the statements provided. The significant differences in the group response means can be attributed to the students' feeling more strongly about many of the items. For example, this result occurred for the statement, "evaluations reflect how strict or lenient a professor grades" ( $t = 5.42, p < .001$ ), and for the statement "the more demanding the professor, the lower the rating" ( $t = 1.97, p < .05$ ). In one instance, however, the two groups felt quite differently about a statement. While students slightly agreed that evaluation results should be published, faculty members disagreed somewhat that the information should be published ( $t = 7.97, p < .001$ ).

#### Discussion

Although it is difficult to generalize from a single study, the results of this research suggest that students and faculty members do differ somewhat in the importance assigned to some teacher traits. While the differences found are of statistical significance, it is not clear that these differences are of practical significance for the groups under investigation. The calculation of the Spearman rank correlation indicates that there are areas of agreement in

each groups' assessment of teacher traits. Reinforcing evidence comes from the use of a formula for determining effect size (or a point-biserial  $r$ ) which indicates, on average, that less than 10 percent of the variance is accounted for by group difference. One might note, however, that there are traits which are rated in a quite divergent manner. These latter differences suggest that care must be taken rating teachers on the basis of student evaluations.

An important area of concern is the research evidence available which identifies a number of factors which are not controllable by teachers yet influence the student ratings of teacher performance.

Faculty members' concern extends beyond the issue of whether these measures are affected by student attributes or other factors beyond their control. Many believe that students give lower ratings to more demanding professors as well as those who are harder graders. Students in the present study disagreed with both of these perceptions, but faculty members were undecided about the latter perception and disagreed with the former statement. These findings might indicate that certain faculty perceptions of how students respond on evaluations are inaccurate. There is agreement by both groups that student evaluations should be used. This would indicate that such ratings have value, if for no other purpose than for providing feedback.

An important issue of concern to faculty members is how the evaluation results should be used. One surprising finding of the present study was agreement that student evaluations should not be

used for personnel decisions. This finding is made even more surprising by the direction of the difference, with students feeling more strongly than faculty that these evaluations should not be used for that purpose ( $t = 6.83$ ,  $p < .001$ ). Finally, one of the central issues imbedded in this study was the question of whether students and faculty use different criteria in evaluating teacher effectiveness. On this issue statistically significant differences were found; however, both groups disagreed with the statement that faculty and students use similar evaluation criteria. Faculty disagreed more strongly ( $t = 6.41$ ).

#### Conclusions

Differences do exist in student and faculty perceptions of both important teacher traits and behaviors and the value and appropriate use of student evaluations. Given these statistical differences, it seems appropriate to exercise caution in using the results of students in making major personnel decisions. The trend is clear, that is, personnel decisions are being made on the basis of such evaluations and that the results of these evaluations are being widely disseminated. In the past, they were used more for development purposes and the results were placed in the hands of the teacher. Until more general agreement is obtained in defining teaching effectiveness and in determining how to better measure this elusive concept, it may be best to return to the stage where

such instruments are used exclusively for the purpose of giving teachers feedback, or only a small weight is given to this information when used for administrative decisions.

Where might we go from here? Future areas of research would include the use of multiple evaluations: self-ratings, peer ratings, student ratings, and supervisor ratings. Classroom visitations, although problematic and often objected to by many teachers, should become a more integral part of this process. Recent articles suggest that an emerging issue is that of confidentiality (e.g., Miller 1978). Until more reliable, valid and useful instruments are developed, results should be placed only in the hands of those who need to know and use the information. In addition, from a research perspective, a more systematic approach needs to be taken in identifying the relevant variables which influence student ratings, and developing as well as testing better instruments. There is a body of literature on performance evaluations that may be useful in identifying relevant variables. This literature would also give researchers a base from which to work when attempting to understand the evaluation process as it occurs in academia. This literature could aid administrators and others vested with the responsibility of performance evaluation by helping them systematically perceive the entire evaluation process from the beginning phase (position analysis) to the final phase (evaluation of personnel by raters utilizing job-related and legally defensible instruments

which have reliability and validity). Academic institutions which choose to utilize teacher evaluations as criteria for administrative decisions must consider not only the reliability and validity issues most often raised, but also the legal ramifications of doing so.

Since student evaluations are being used as input for personnel decisions, it seems appropriate to view these instruments as selection devices. As such, these devices are subject to close scrutiny by the Equal Employment Opportunity Commission as well as similar regulatory agencies at other levels. The research previously reported indicates that student evaluations would fail to meet the reliability and validity standards imposed by such regulatory bodies. Although many student evaluations are in use today, only a few have undergone the extensive refinement and rigorous validation process that personnel psychologists and psychometricians have utilized in sectors outside of academe. The question that comes to mind is: do we wait until professors who feel wronged by the student evaluation process take legal action, or do we anticipate such challenges by taking whatever steps are necessary to develop instruments which are reliable and valid?

Reference Notes

1. Baba, V., & Ace, M. E. Leader behavior in the university classroom. Paper presented at the meeting of the Western American Institute for Decision Sciences, San Diego, March, 1978.
2. Banziger, G. J., & Smith, R. K. Student motivation and course evaluation. Paper presented at the meeting of the Western American Institute for Decision Sciences, San Diego, March, 1978.
3. Baum, P. A., & Brown, W. W. Student and faculty perceptions of teaching effectiveness. Paper presented at the meeting of the Western American Institute for Decision Sciences, San Diego, March, 1978.
4. Weaver-Wilson, P., & Wilson, T. C. Differences in student evaluations from business vs. other colleges. Paper presented at the meeting of the Southern Marketing Association, Atlanta, October, 1977.

References

- Centra, J. A. College teaching: Who should evaluate it? Findings, 1974, 1, 5-8.
- Costin, F., Greenough, W. T., & Menges, R. J. Student ratings of college teaching: Reliability, validity, and usefulness. Review of Education Research, 1971, 41, 511-535.
- Cummings, L. L., & Schwab, D. P. Performance in organizations: Determinants and appraisal. Glenview, Illinois: Scott, Foresman, 1973, 106.
- Elmore, P. B., & LaPointe, K. A. Effect of teacher sex, student sex and teacher warmth on the evaluation of college instructors. Journal of Educational Psychology, 1975, 67, 368-374.
- Kinicki, A. & Schriesheim, C. A. Teachers as leaders: A moderator variable approach. Journal of Educational Psychology, 1978, 70, 928-935.
- Lein, C. D. & Merz, M. Faculty evaluation in schools of business: The impact of AACSB accreditation on promotion and tenure decisions. Collegiate News and Views, 1978, 31 (2), 21-24.
- Maier, N. R. F., Hoffman, L. R., Hooven, J. J., & Kead, W. H. Superior-subordinate communication: A statistical research project, New York: American Management Association, 1961.
- Miller, D. V. Dangers of using student evaluations for administrative purposes. Collegiate News and Views, 1978, 31 (3), 2-4.

- Miller, D. W., Brokaw, S., & Shaaban, F. Administrative uses of student evaluation of college instruction. Proceedings of the Annual Meeting of the American Institute for Decision Sciences, 1977, 589.
- Penfield, D. A. Student ratings of college teaching: Rating the utility of rating forms. The Journal of Educational Research, 1978, 72 (1), 19-22.
- Peterson, R. A., Kerin, R. A., & Martin, W. S. Teaching effectiveness in the decision sciences: The use of student evaluations. Decision Sciences, 1978, 9, 712-714.
- Rodin, M. & Rodin, B. Student evaluation of teachers. Science, 1972, 177, 1164-1166.
- Schwab, D. P. Attitudes measured prior to taking a course and subsequent evaluations of that course and instructor. Proceedings of the Annual Meeting of the Academy of Management Proceedings, 1977, 454.
- Sullivan, A. M., & Skanes, G. R. Validity of student evaluation of teaching and the characteristics of successful instructors. Journal of Educational Psychology, 1974, 66, 584-590.
- Swanson, R. G. Teacher evaluation as a function of leadership style: A multiple correlational approach. Unpublished doctoral dissertation, North Texas State University, 1974.
- Warren, W. & O'Connell, W. E. The effect of instructor reputation and personality upon instructor evaluations. Proceedings of the Annual Meeting of the Academy of Management, 1977, 45-49.

Wilson, D., & Doyle, K. O. Student ratings of instruction.

Journal of Higher Education, 1976, 47, 465-470.

Witty, P. An analysis of the personality traits of the effective teacher. Journal of Educational Research, 1947, 40, 662-671.

Table 1

## SELECTED TEACHER TRAITS

| Teaching Trait                           | Faculty (N=217) |      |      | Students (N=559) |      |      | t-value |            |
|------------------------------------------|-----------------|------|------|------------------|------|------|---------|------------|
|                                          | Rank            | Mean | S.D. | Rank             | Mean | S.D. | Mean    | Difference |
| 1. Lectures are easy to outline          | 7               | 5.08 | 1.30 | 5                | 5.61 | 1.41 | 4.22    | **         |
| 2. Lectures are entertaining             | 10              | 4.61 | 1.54 | 8                | 5.22 | 1.48 | 4.44    | **         |
| 3. Indicates what is important for exams | 8               | 5.03 | 1.43 | 2                | 6.06 | 1.20 | 8.87    | **         |
| 4. Expects students to be prepared       | 4               | 5.95 | 1.11 | 6                | 5.50 | 1.29 | 3.92    | **         |
| 5. Emphasizes factual knowledge          | 9               | 4.78 | 1.46 | 7                | 5.40 | 1.36 | 4.83    | **         |
| 6. Emphasizes concepts                   | 1               | 6.41 | .85  | 1                | 6.12 | 1.02 | 3.35    | **         |
| 7. Stresses applications                 | 2               | 6.28 | .88  | 3                | 6.03 | 1.13 | 2.55    | *          |
| 8. High grading standards                | 6               | 5.15 | 1.43 | 9t               | 4.64 | 1.36 | 4.04    | **         |
| 9. Creative thinking on exams            | 5               | 5.38 | 1.25 | 9t               | 4.64 | 1.60 | 5.37    | **         |
| 10. Exhibits concern for students        | 3               | 5.98 | 1.22 | 4                | 5.84 | 1.31 | 1.19    |            |

NOTE: Group means range from 7 = extremely important to 1 = not at all important

\*p < .05

\*\*p < .001

Table 2

| STUDENT EVALUATION RESPONSES                                           |                 |      |                  |      |                 |
|------------------------------------------------------------------------|-----------------|------|------------------|------|-----------------|
| Statement                                                              | Faculty (N=217) |      | Students (N=559) |      | t-value         |
|                                                                        | Mean            | S.D. | Mean             | S.D. | Mean Difference |
| 1. Evaluations measure teaching effectiveness                          | 3.86            | 1.70 | 4.25             | 1.66 | 2.88**          |
| 2. Students use evaluations to punish professors                       | 4.18            | 1.70 | 3.27             | 1.87 | 6.25***         |
| 3. Students are qualified to judge professors                          | 4.35            | 1.61 | 4.87             | 1.64 | 3.98***         |
| 4. Evaluations serve no useful function                                | 2.29            | 1.50 | 2.71             | 1.71 | 3.20***         |
| 5. Evaluations should not be used                                      | 2.83            | 1.68 | 2.73             | 1.64 | <1              |
| 6. Only professors have the knowledge to rate professors               | 2.20            | 1.45 | 1.91             | 1.34 | 2.66**          |
| 7. Student evaluations should be discontinued                          | 2.25            | 1.61 | 2.22             | 1.50 | <1              |
| 8. Evaluation results should be published                              | 3.01            | 1.98 | 4.37             | 2.19 | 7.97***         |
| 9. Evaluations are an important means of measuring performance         | 4.48            | 1.69 | 4.81             | 1.55 | 2.64**          |
| 10. The more demanding the professor, the lower the rating             | 3.39            | 1.76 | 3.11             | 1.74 | 1.97*           |
| 11. Evaluations reflect how hard or how easy a professor grades        | 3.90            | 1.70 | 3.17             | 1.68 | 5.42***         |
| 12. All other performance measures are superior to teacher evaluations | 2.36            | 1.30 | 2.79             | 1.39 | 3.55***         |
| 13. Teacher evaluations should not be used for personnel decisions     | 3.97            | 1.88 | 5.03             | 1.95 | 6.83***         |
| 14. Once completed, evaluations are discarded and never seen again     | 2.65            | 1.62 | 3.47             | 1.75 | 5.99***         |
| 15. Students and faculty use the same criteria to evaluate performance | 2.60            | 1.43 | 3.42             | 1.67 | 6.41***         |

\*p &lt; .05

\*\*p &lt; .01

\*\*\*p &lt; .001

NOTE: Group means range from 7 = strongly agree, 1 = strongly disagree